

Coverage- and Nonresponse-Effekte bei Online-Bevölkerungsumfragen

Bandilla, Wolfgang; Kaczmirek, Lars; Blohm, Michael; Neubarth, Wolfgang

Veröffentlichungsversion / Published Version
Sammelwerksbeitrag / collection article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Bandilla, W., Kaczmirek, L., Blohm, M., & Neubarth, W. (2009). Coverage- and Nonresponse-Effekte bei Online-Bevölkerungsumfragen. In N. Jakob, H. Schoen, & T. Zerback (Hrsg.), *Sozialforschung im Internet: Methodologie und Praxis der Online-Befragung* (S. 129-144). Wiesbaden: VS Verl. für Sozialwiss. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-46471-2>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public. By using this particular document, you accept the above-stated conditions of use.

Coverage- und Nonresponse-Effekte bei Online-Bevölkerungsumfragen

Von Wolfgang Bandilla, Lars Kaczmarek, Michael Blohm & Wolfgang Neubarth

1. Einleitung und Hintergrund¹

Kaum eine Technologie hat sich so schnell verbreitet und das mediale und kommunikative Verhalten weiter Bevölkerungskreise verändert wie das Internet. Es kann deshalb auch nicht überraschen, dass die Möglichkeiten dieser Technik innerhalb der Markt- und Sozialforschung zunehmend an Bedeutung gewinnen. Die auf das Jahr 2006 bezogenen Zahlen des *Arbeitskreises Deutscher Markt- und Sozialforschungsinstitute e.V.* (ADM)² weisen z.B. aus, dass beim Vergleich der gängigen Befragungsarten der Anteil von Online-Interviews auf mittlerweile 21 Prozent gestiegen ist und damit weit über dem traditionell-schriftlicher Befragungen liegt (acht Prozent). Der Anteil persönlich-mündlicher Interviews liegt mit 25 Prozent nur geringfügig höher, während telefonische Befragungen mit einem Anteil von 46 Prozent deutlich dominieren. Auch wenn nach wie vor Befragungen überwiegend interviewer-administriert (d.h. telefonisch oder persönlich-mündlich) durchgeführt werden, ist der kurze Zeitraum bemerkenswert, in dem die Online-Erhebung eine nennenswerte Größenordnung erreicht hat: Lag ihr Anteil im Jahr 2002 bei lediglich fünf Prozent, so hat er sich innerhalb von nur vier Jahren auf über 20 Prozent erhöht.

Ein Grund für diese Entwicklung dürfte in spezifischen Vorteilen der Online-Erhebung liegen: Als selbstadministrierte Erhebungsart bietet sie im Vergleich zu den interviewer-administrierten Befragungen deutliche Kostenvorteile. Hinzu kommen Zeitvorteile (d.h. vergleichsweise kurze Feldzeiten) und die Möglichkeit, den Befragungsprozess interaktiv zu gestalten.³ Diese Vorteile werden gegenwärtig vor allem bei Studien im Bereich der Marktforschung genutzt, also einem Bereich, in dem bevölkerungsrepräsentative Aspekte keine oder nur eine untergeordnete Rolle spielen.

Eine grundlegend andere Situation ist jedoch gegeben, wenn die Ergebnisse einer Umfrage auf die Allgemeinbevölkerung generalisiert werden sollen. Nach wie vor verfügen weite Bevölkerungskreise über keinen Zugang zum Internet. Zwar nutzen nach den aktuellen Ergebnissen der ARD/ZDF-Online-Studie⁴ 63 Prozent aller Personen ab 14 Jahren zumindest gelegentlich das Internet, was umgekehrt aber auch bedeutet, dass nahezu 40 Prozent nicht an Online-Befragungen teilnehmen können. Trotz der Dynamik, die bei der Verbreitung dieses neuen Mediums innerhalb weniger Jahre zu verzeichnen war (Anstieg zwischen 1997 und 2007 von sechs Prozent auf 63 Prozent), unterscheiden sich die Internetnutzer weiterhin in zentralen soziodemographischen Variablen wie Alter, Geschlecht und Bildungsabschluss von denjenigen, die über dieses Medium nicht erreicht werden können.

¹ Unser Dank gilt Mick Couper für wertvolle Anregungen und Kommentare.

² Vgl. ADM 2007.

³ Vgl. u.a. Reips & Bosnjak 2001; Dillman 2007.

⁴ Vgl. van Eimeren & Frees 2007: 363f.

Anhand des Merkmals Alter lassen sich die unterschiedlichen Realisierungsmöglichkeiten bzw. Grenzen von Online-Befragungen beispielhaft verdeutlichen. Sowohl die ARD/ZDF-Studie als auch andere Studien⁵ kommen zu dem Ergebnis, dass bei der Altersgruppe der bis zu 30-Jährigen eine nahezu vollständige Internetnutzung gegeben ist. D.h. bei dieser Altersgruppe ist die Voraussetzung einer prinzipiellen Teilnahme an einer Online-Befragung gegeben – systematische Ausschlüsse von Befragten aufgrund eines fehlenden Zugangs zum Internet können bei dieser Altersgruppe weitgehend vernachlässigt werden (lässt man die Probleme der Stichprobenziehung einmal beiseite). Ein Beispiel für die erfolgreiche Realisierung einer Online-Befragung innerhalb dieser Altersgruppe stellt der HISBUS dar. Es handelt sich hierbei um eine repräsentative Befragung von Studierenden an deutschen Hochschulen, die vom HIS Hochschul-Informationssystem durchgeführt wird.⁶ Demgegenüber werden die (gegenwärtigen) Grenzen der Online-Erhebung sichtbar, wenn aus inhaltlichen Erkenntnisinteressen ältere Bevölkerungsgruppen im Fokus einer Befragung stehen. Laut Angaben der ARD/ZDF-Studie umfasst z.B. die Altersgruppe der ab 60-Jährigen in absoluten Zahlen ca. 20,4 Millionen Personen, von denen mit 25 Prozent lediglich ein Viertel das Internet nutzt. D.h. bei einer Online-Befragung wären drei Viertel dieser Altersgruppe (also ca. 15 Millionen Personen) systematisch von der Teilnahme ausgeschlossen.

Die Ergebnisse einer jüngst veröffentlichten amerikanischen Studie⁷ belegen das dabei auftretende Problem: In dieser Studie wurde eine Zufallsstichprobe von Personen im Alter von 50 Jahren und älter entweder (1) traditionell oder (2) online zu gesundheitsbezogenen Themen befragt. Die Online-Erhebungsart war nur bei denjenigen Personen möglich, die über einen Internetzugang verfügten (ca. 30 Prozent der Gesamtstichprobe). Beim Vergleich der beiden Gruppen zeigen sich z.T. deutliche Unterschiede in den Ergebnissen, so dass sich allein auf der Basis der online erhobenen Daten nicht valide auf die Gesamtheit dieser Altersgruppe schließen lässt.

Anhand der einschlägigen Methodenliteratur⁸ lassen sich die im zweiten Beispiel aufgezeigten Unterschiede mit „Undercoverage“ erklären: Die angestrebte Grundgesamtheit bilden alle Personen im Alter 50+. Die Auswahlgesamtheit würden bei einer ausschließlich online durchgeführten Befragung nur diejenigen Personen darstellen, die über einen Zugang zum Internet verfügen (im vorliegenden Beispiel 30 Prozent). Der Fehler, der sich aus der Differenz von 70 Prozentpunkten zwischen angestrebter Grundgesamtheit und Auswahlgesamtheit ergibt, ist bekannt als Undercoverage Error. Die bei Umfragen zugrundeliegenden Stichprobendesigns haben jedoch üblicherweise das Ziel, diesen Fehler möglichst gering zu halten, d.h. eine große Übereinstimmung zwischen Grundgesamtheit und Auswahlgesamtheit anzustreben.⁹

Die hier dargestellte Problematik speziell bei Online-Befragungen in der Allgemeinbevölkerung ist nicht neu¹⁰, und zur Lösung werden verschiedene Möglichkeiten vorgeschlagen, die im Folgenden kurz erwähnt werden. Die radikalste (jedoch auch sehr kostenintensive) Lösung besteht darin, allen Personen einer Zufallsstichprobe, die über keinen Zugang zum Internet verfügen, die für die Teilnahme an Online-Befragungen notwendige techni-

⁵ Vgl. Medienpädagogischer Forschungsverbund Südwest 2007.

⁶ Vgl. Müßig-Trapp et al. 2006.

⁷ Vgl. Couper et al. 2007.

⁸ Vgl. u.a. Groves et al. 2004.

⁹ Vgl. hierzu auch Schnell et al. 2005: 272.

¹⁰ Vgl. u.a. Couper 2000; Faas 2003.

sche Ausstattung zur Verfügung zu stellen. Dieser Ansatz wird seit einigen Jahren in den USA von Knowledge Networks¹¹ kommerziell umgesetzt. Eher unter akademischer Forschungsperspektive stehend startete das MESS-Projekt, das an der Universität Tilburg in den Niederlanden angesiedelt ist. Die Projekt-Website¹² gibt dabei Auskunft über die zur Realisierung dieses Projektes zur Verfügung stehenden finanziellen Mittel (13,9 Millionen Euro). Alternativ zu diesen sehr aufwändigen Vorhaben werden Möglichkeiten diskutiert, den Coverage Error durch eine Gewichtung der online erhobenen Daten zu reduzieren. Im Mittelpunkt stehen dabei vor allem „Propensity Score Adjustments (PSA)“.¹³ Die Ergebnisse derartiger Modellierungen sind nicht immer eindeutig und erfordern für eine umfassende Bewertung sicherlich noch weitere Forschungsaktivitäten. Hierzulande am bekanntesten ist das von der Unternehmensberatung McKinsey finanzierte Online-Umfrageprojekt „Perspektive Deutschland“, bei dem eine solche Gewichtung bzw. Modellierung umgesetzt wurde.¹⁴

Gleichgültig welche gegenwärtigen Lösungsmöglichkeiten gesehen werden, es bleibt festzuhalten, dass die unterschiedlichen Nutzungen bzw. Online-Erreichbarkeiten und der damit verbundene Coverage Error ein gewichtiges Problem darstellen. Hierauf dürfte sich auch die nach wie vor und insbesondere in der akademisch ausgerichteten Umfrageforschung weit verbreitete Skepsis gegenüber Online-Bevölkerungsumfragen zurückführen lassen. Eine Situation, die Parallelen zu derjenigen aufweist, die zu Beginn telefonischer Interviews bestand.¹⁵

Diese Skepsis wird zusätzlich dadurch gefördert, dass Online-Befragungen (ähnlich wie schriftliche Befragungen) vergleichsweise geringe Teilnahmequoten¹⁶ aufweisen.¹⁷ D.h. neben dem Coverage Error ist (bei gegebenem Internetzugang) als potenziell weitere Fehlerquelle der Nonresponse-Error zu beachten, also der Fehler, der sich aus der Differenz zwischen den Angaben der Befragungsteilnehmer und der Befragtengesamtheit ergibt.

Mit der nachfolgend vorgestellten Studie wird der Versuch unternommen, diese beiden Fehler anhand adäquater empirischer Daten zu quantifizieren und somit eine bessere Einschätzung von Coverage- und Nonresponse-Effekten bei Online-Befragungen zu ermöglichen. Die Grundlage bilden dabei Daten der im Jahr 2006 durchgeführten Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS).

¹¹ Vgl. u.a. Dennis & Li 2003.

¹² Vgl. <http://www.uvt.nl/centerdata/en/mess/>.

¹³ Vgl. Lee 2006.

¹⁴ Vgl. Fries 2006; <http://www.perspektive-deutschland.de>.

¹⁵ Vgl. Frey, Kunz & Lüschen 1990.

¹⁶ Gemeint sind hier einmalige Online-Befragungen (Querschnitt-Befragungen) und nicht Online-Access-Panels, bei denen eine grundsätzlich anders gelagerte Ausgangssituation vorliegt.

¹⁷ Vgl. u.a. Schonlau et al. 2002.

2. Methode und Untersuchungsanlage

Die Berechnung der Coverage- und Nonresponse-Effekte setzte eine Studienkonzeption voraus, die es ermöglichte, zwischen folgenden Gruppen unterscheiden zu können: a) Internetnutzer vs. Nicht-Nutzer des Internet, b) zur Teilnahme an einer Online-Befragung bereit vs. nicht bereit Personen (bei vorhandenem Internetzugang) und c) Teilnehmer vs. Nicht-Teilnehmer einer Online-Befragung (bei vorhandener Teilnahmebereitschaft). Um dieses Ziel zu erreichen, erfolgte die Teilnehmerrekrutierung der hier beschriebenen Studie über den ALLBUS.

Beim ALLBUS handelt es sich um eine Umfrageserie zu Einstellungen, Verhaltensweisen und zur Sozialstruktur der Bevölkerung in der Bundesrepublik Deutschland, die seit 1980 in zweijährigem Abstand durchgeführt wird.¹⁸ In persönlichen Interviews wird jeweils ein repräsentativer Querschnitt der bundesdeutschen Bevölkerung zu wechselnden Themenschwerpunkten befragt. Die Grundgesamtheit besteht dabei aus der erwachsenen Wohnbevölkerung (d.h. Deutschen und deutschsprachigen Ausländern in West- und Ostdeutschland). Als Serviceleistung für die sozialwissenschaftliche Forschung und Lehre werden die Daten allen interessierten Personen und Institutionen für Analysen zur Verfügung gestellt.

Bei der im Jahr 2006 durchgeführten ALLBUS-Erhebung, welche die Grundlage der hier vorgestellten Studie bildet, wurde die Stichprobe mit Hilfe eines zweistufigen Ziehungsverfahrens gebildet: Zunächst wurde eine Stichprobe von Gemeinden gezogen und anschließend wurden in den Gemeinden aus den Einwohnermelderegistern Personenadressen ausgewählt.¹⁹ Befragt wurden insgesamt 3.421 Personen, davon 2.299 in West- und 1.122 in Ostdeutschland. D.h. es wurde ein disproportionaler Stichprobenansatz mit einem Oversampling in Ostdeutschland gewählt, um auch für die neuen Bundesländer eine ausreichende Fallzahl für differenzierte Analysen zur Verfügung zu stellen. Die Ausschöpfungsquote lag bei 40,2 Prozent in West- und 42,8 Prozent in Ostdeutschland. Zur Prüfung der Stichprobenqualität wurden die Verteilungen standarddemographischer Variablen der ALLBUS-Nettostichprobe mit den entsprechenden Verteilungen des Mikrozensus abgeglichen. Mit Ausnahme des bekannten Bildungs-Bias (Mittelschicht-Bias)²⁰ wurden keine gravierenden Abweichungen festgestellt.

Der thematische Schwerpunkt dieser ALLBUS-Umfrage bestand in „Einstellungen gegenüber ethnischen Gruppen“. Zwei im Kontext dieses Schwerpunktes stehende Fragebatterien bildeten die Basis für die Analyse des Nonresponse Errors: Zum einen eine Frage, mit der anhand von vier Items „Einstellungen gegenüber in Deutschland lebenden Ausländern“ erhoben wurden. Zum anderen eine Frage, in der acht Items die „wahrgenommenen Konsequenzen der Anwesenheit von Ausländern“ erfassen.²¹

Die für diese Studie zentralen Fragen zur Internetnutzung und zur Bereitschaft, an einer Online-Befragung teilzunehmen, wurden am Ende der durchschnittlich 50 Minuten dauernden persönlichen Interviews gestellt. Die erste Frage bezog sich auf die private Internetnut-

¹⁸ Umfassende Informationen zum ALLBUS finden sich auf der Website der GESIS unter <http://www.gesis.org/Dauerbeobachtung/Allbus/>.

¹⁹ Ausführlich hierzu Wasmer et al. 2007: 52.

²⁰ Gemeint ist hier der Anteil von Personen mit niedrigen Bildungsabschlüssen, der bei derartigen Umfragen erfahrungsgemäß unter dem betreffenden Anteil des Mikrozensus liegt und bei dieser Umfrage ein mittleres Niveau aufweist. Vgl. hierzu Wasmer et al. 2007: 73f.

²¹ Der Wortlaut dieser Fragen findet sich im Anhang.

zung und lautete: „Nutzen Sie privat das Internet?“ Die Einschränkung auf die private Nutzung wurde bewusst gewählt, da wir davon ausgehen, dass die Teilnahme an derartigen Befragungen nur im privaten Umfeld und nicht z.B. während der Arbeitszeit erfolgen kann.²² Sofern die Frage nach der Nutzung bejaht wurde, folgte eine kurze Frage nach der Nutzungshäufigkeit und die entscheidende Rekrutierungsfrage: „Wären Sie bereit, unser Forschungsprojekt auch weiterhin durch Ihre Mitarbeit bei einer Internet-Umfrage zu unterstützen?“ Wurde auch diese Frage bejaht, erfolgte die Aufnahme der E-Mail-Adresse durch den Interviewer. Nach Abschluss der ALLBUS-Feldphase wurden alle befragungsbereiten Personen mittels dieser E-Mail-Adressen eingeladen, an der Online-Befragung teilzunehmen.

Die hier beschriebene Vorgehensweise ermöglicht die Bildung der schon weiter oben beschriebenen Vergleichsgruppen und damit die Berechnung von Coverage und Nonresponse Error anhand der in den ALLBUS-Interviews erhobenen Daten. Der Vorteil dieser Untersuchungsanlage besteht darin, über den gesamten Rekrutierungsprozess hinweg das Ausmaß der Fehler direkt ablesen zu können, da sowohl Daten aus der jeweils interessierenden Gesamtheit, wie auch den Antwortenden und den Nicht-Antwortenden vorliegen. Bspw. errechnet sich der Coverage Error aus der Differenz der Mittelwerte der interessierenden Variablen zwischen Internetnutzern und Gesamtheit (hier: ALLBUS-Stichprobe). Gleiches gilt für den Nonresponse Error, z.B. bei gegebener Teilnahmebereitschaft die Differenz zwischen Antwortenden und der Gesamtheit teilnahmebereiter Personen.²³

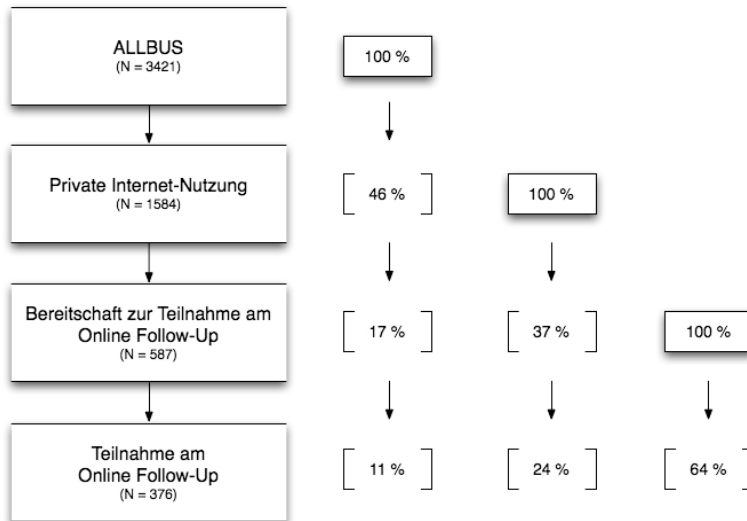
²² Ähnlich wie bei telefonischen oder persönlichen Interviews, die ausnahmslos über Privatanschlüsse bzw. in den Privathaushalten (und nicht am Arbeitsplatz!) erfolgen.

²³ Zur Berechnung von Coverage und Nonresponse Error vgl. Groves et al. 2004: 54ff.

3. Ergebnisse²⁴

Abbildung 1 gibt eine Übersicht zu den Fallzahlen, die im Rekrutierungsverlauf für eine Online-Befragung auftreten, sofern diese – wie im vorliegenden Fall – über persönliche und auf einer Zufallsstichprobe basierende Interviews erfolgt.

Abbildung 1: Rekrutierungsverlauf



Auf den ersten Blick überraschend ist der mit 46 Prozent vergleichsweise geringe Anteil derjenigen, die zu privaten Zwecken das Internet nutzen. Andere Umfragen, wie z.B. die schon erwähnte ARD/ZDF-Studie kommen bzgl. der „gelegentlichen Onlinenutzung“²⁵ auf deutlich über 60 Prozent. Die Differenz dürfte zum einen auf die unterschiedlichen Auswahlgesamtheiten in beiden Studien zurückzuführen sein (ARD/ZDF: 14 Jahre und älter; ALLBUS: 18 Jahre und älter), zum anderen – und weitaus bedeutsamer für die Rekrutierung von Online-Befragungsteilnehmern – auf die Formulierung der Nutzungsabfrage: Bei der ARD/ZDF-Studie (deren Ziel nicht in der Rekrutierung von Befragungsteilnehmern besteht) wird allgemein und unabhängig von der Nutzungsart nach der Online-Nutzung gefragt. Beim ALLBUS wurde dagegen bewusst eine deutlich restriktivere Formulierung gewählt, bei der die Internutzung z.B. während der Arbeitszeit ausgeschlossen ist (und damit für Bevölkerungsumfragen, ähnlich wie in anderen Erhebungsmodi, geeigneter ist).

Des Weiteren überrascht der mit 37 Prozent geringe Anteil Befragungsbereiter bei vorhandenem Zugang zum Internet.²⁶ In der Frage nach der Bereitschaft zur Online-Teilnahme

²⁴ Alle Ergebnisdarstellungen und Analysen berücksichtigen das personenbezogene West-Ost-Gewicht, um das Oversampling in Ostdeutschland auszugleichen.

²⁵ Vgl. van Eimeren & Frees 2007: 364.

²⁶ Hier spielt sicherlich die Erfahrung im Umgang mit dem Medium Internet eine Rolle, was sich z.B. darin zeigt, dass 54 Prozent derjenigen, die sich zur Teilnahme am Online Follow-Up bereit erklärten, täglich das Internet nutzen. Unter den nicht Teilnahmebereiten liegt der Vergleichswert bei 25 Prozent.

liegt offensichtlich eine weitere entscheidende Hürde im Rekrutierungsprozess, die – sofern sie erfolgreich genommen wurde – zu einer mit 64 Prozent hohen Teilnahmequote führt. Die im Rekrutierungsverlauf auftretenden Selektionsmechanismen kommen insbesondere beim Vergleich soziodemographischer Hintergrundmerkmale zum Ausdruck (vgl. Tab. 1).

Tabelle 1: Vergleich soziodemographischer Hintergrundmerkmale

	Schulbildung						Summe Geschlecht %	Summe %
	einfach		mittel		hoch			
	18-39	40+	18-39	40+	18-39	40+		
	%	%	%	%	%	%		
ALLBUS (n = 3309)*								
Männer	5	15	5	9	5	9	48	
Frauen	3	17	7	12	6	7	52	
Summe Schulbildung/Alter	8	32	12	21	11	16		
Summe Schulbildung		40		33		27		100
Summe							100	
Internetnutzung (n = 1553)*								
Männer	5	8	7	9	10	13	52	
Frauen	3	5	11	10	10	9	48	
Summe Schulbildung/Alter	8	13	18	19	20	22		
Summe Schulbildung		21		37		42		100
Summe							100	

* Abweichungen zu den in Abbildung 1 aufgeführten Fallzahlen erklären sich durch fehlende Werte auf einzelnen Kovariaten.

Tabelle 1: Vergleich soziodemographischer Hintergrundmerkmale (Fortsetzung)

	Schulbildung						Summe Geschlecht %	Summe %
	einfach		mittel		hoch			
	18-39	40+	18-39	40+	18-39	40+		
	%	%	%	%	%	%		
Bereitschaft zur Online-Teilnahme (n = 576)*								
Männer	4	6	7	9	14	15	55	
Frauen	2	4	9	8	14	8	45	
Summe Schulbildung/Alter	6	10	16	17	28	23		
Summe Schulbildung		16		33		51		100
Summe							100	
Online-Teilnahme (n = 372)*								
Männer	2	5	5	11	15	16	54	
Frauen	2	4	10	8	14	8	46	
Summe Schulbildung/Alter	4	9	15	19	29	24		
Summe Schulbildung		13		34		53		100
Summe							100	

* Abweichungen zu den in Abbildung 1 aufgeführten Fallzahlen erklären sich durch fehlende Werte auf einzelnen Kovariaten.

Die deutlichsten Unterschiede zwischen der ALLBUS-Gesamtgruppe und der Subgruppe der Internetnutzer zeigen sich bzgl. der Merkmale Alter und Schulabschluss²⁷: Bezogen auf alle ALLBUS Befragten liegt der Anteil derjenigen, die über einen einfachen Schulabschluss verfügen und 40 Jahre und älter sind, bei 32 Prozent. Unter den Internetnutzern sind es lediglich 13 Prozent. Ein vollkommen gegensätzliches Bild zeigt sich bei der Alters-

²⁷ Der entsprechenden Variable liegt die Frage nach dem allgemeinen Schulabschluss zugrunde: einfache Schulbildung (Volks-/Hauptschulabschluss), mittel (Realschulabschluss), hoch (Fachhochschulreife, Abitur). Nicht eindeutige Angaben (z.B. bei „anderer Schulabschluss“) wurden nicht berücksichtigt („missing“).

gruppe der 18- bis 39-Jährigen mit einem hohen Bildungsabschluss: Hier liegt der Anteil unter den Internetnutzern mit 20 Prozent nahezu doppelt so hoch im Vergleich zur entsprechenden Gruppe beim Gesamt-ALLBUS (11 Prozent). Diese Unterschiede sind bereits aus früheren Studien²⁸ bekannt und belegen auf der soziodemographischen Ebene den „Bias“ (Verzerrung), der derzeit bei einer ausschließlich online durchgeführten Bevölkerungsbefragung auftreten würde.

Aber auch dann, wenn über einen Internetzugang die technischen Voraussetzungen zur Teilnahme an einer Online-Befragung gegeben sind, ist ein deutlicher Bias bei den Befragungsbereiten zu erkennen: Innerhalb dieser Gruppe sind im Vergleich zur Gesamtgruppe der Internetnutzer wiederum die „Jüngeren“ mit einem hohen Bildungsabschluss überrepräsentiert (28 Prozent vs. 20 Prozent). Wird dagegen auf der letzten Rekrutierungsstufe die Bereitschaft zur Online-Teilnahme gegeben, so sind zumindest auf der soziodemographischen Ebene nur marginale Unterschiede zu den tatsächlich an der Online-Umfrage Teilnehmenden zu erkennen (28 bzw. 29 Prozent).

D.h. bei der Rekrutierung von Teilnehmern für eine Online-Befragung ist – zumindest nach diesen Ergebnissen – neben dem Coverage Error (fehlender Internetzugang) auch der hohe Anteil derjenigen zu beachten, die bei vorhandenem Zugang zum Internet nicht zur Online-Teilnahme bereit sind („Nonrespondents“). Zusammengefasst haben Jüngere mit hohem Bildungsabschluss eine erhöhte Wahrscheinlichkeit, die einzelnen (zwingend vorhandenen) Auswahlstufen bis zur tatsächlichen Teilnahme zu durchlaufen.

Wie wirken sich nun diese Unterschiede zwischen den verschiedenen Gruppen auf den Coverage bzw. Nonresponse Error substantieller Variablen aus? Die Beantwortung dieser Frage soll auf Grundlage zweier zentraler Fragebatterien zum Themenschwerpunkt des ALLBUS 2006 erfolgen. Erhoben wurden dabei das Ausmaß der Zustimmung zu „Einstellungen gegenüber in Deutschland lebenden Ausländern“ (vier Items) und die „wahrgenommenen Konsequenzen der Anwesenheit von Ausländern“ (acht Items). Die Beantwortung erfolgte über eine 7-stufige Zustimmungsskala mit den Endpunkten „stimme überhaupt nicht zu“ (Wert 1) und „stimme voll und ganz zu“ (Wert 7). Berechnet wurden Mittelwerte und Standardfehler für die jeweiligen Gesamt- und Subgruppen (vgl. Tabellen 2 und 3).

Nahezu durchgängig sind die im Rekrutierungsverlauf entstehenden soziodemographischen Unterschiede sowohl im Coverage als auch im Nonresponse Error zu erkennen. Die vergleichsweise stärkste Abweichung der Item-Mittelwerte ist zwischen Internetnutzern und der ALLBUS-Gesamtgruppe zu verzeichnen (Coverage Error).²⁹ Bei gegebenem Internetzugang ist der durch die fehlende Bereitschaft zur Online-Teilnahme bedingte Nonresponse Error nahezu ausnahmslos größer als der (zweite) Nonresponse Error, der bei gegebener Teilnahmebereitschaft durch die Nicht-Teilnahme verursacht wird. D.h. die wesentliche Hürde hinsichtlich der Generalisierbarkeit online erhobener Daten liegt in der Nicht-Erreichbarkeit derjenigen, die über keinen Internetzugang verfügen. Der bei vorhandenem Internetzugang zusätzlich auftretende Nonresponse Error fällt größtenteils geringer aus.

²⁸ Vgl. u.a. Bandilla 2003.

²⁹ Lediglich bei den Items „E“ und „H“ in Tabelle 3 fällt der Coverage Error geringer aus. Item „A“ in Tabelle 3 zeigt diesbezüglich so gut wie keinen Unterschied.

Tabelle 2: Coverage und Nonresponse Error bei „Einstellungen gegenüber in Deutschland lebenden Ausländern“

	ALLBUS Gesamt (n=3185)*	Internet- nutzung (n=1491)*	Bereit- schaft Teil- nahme (n=560)*	Online- Teil- nahme (n=359)*	Coverage Error	Nonresponse Error		
	(1)	(2)	(3)	(4)	Internet- nutzung (2 zu 1)	Bereit- schaft (3 zu 2)	Teil- nahme (4 zu 3)	Gesamt (4 zu 1)
A) Ausländer: mehr Lebensstilanpassung ³⁰								
<i>M</i>	5.75	5.53	5.38	5.33	-0.22	-0.15	-0.05	-0.42
<i>SE</i>	.027	.041	.068	.083				
B) Ausländer: wieder heim bei knapper Arbeit								
<i>M</i>	3.22	2.78	2.71	2.63	-0.44	-0.07	-0.08	-0.59
<i>SE</i>	.035	.046	.074	.091				
C) Ausländer: polit. Betätigung untersagen								
<i>M</i>	3.33	2.97	2.80	2.67	-0.36	-0.17	-0.13	-0.66
<i>SE</i>	.038	.053	.084	.100				
D) Ausländer: sollten unter sich heiraten								
<i>M</i>	2.46	1.85	1.61	1.57	-0.61	-0.24	-0.04	-0.89
<i>SE</i>	.034	.039	.058	.069				

* Abweichungen zu den in Abbildung 1 aufgeführten Fallzahlen erklären sich durch fehlende Werte auf einzelnen Variablen.

Der Gesamt-Fehler (Coverage plus Nonresponse Error) bewegt sich bei der Mehrzahl der hier analysierten Items im Bereich zwischen 0.4 und 0.6, der Maximalwert liegt bei -0.89 (Item „D“ in Tabelle 2).

³⁰ Vollständiger Wortlaut der Aussagen im Anhang

Tabelle 3: Coverage und Nonresponse Error bei „wahrgenommenen Konsequenzen der Anwesenheit von Ausländern“

	ALLBUS Gesamt (n=3185)	Internet- nutzung (n=1491)	Bereit- schaft Teil- nahme (n=560)	Online- Teil- nahme (n=359)	Coverage Error	Nonresponse Error		
	(1)	(2)	(3)	(4)	Internet- nutzung (2 zu 1)	Bereit- schaft (3 zu 2)	Teil- nahme (4 zu 3)	Gesamt (4 zu 1)
A) Ausländer tun die unschönen Arbeiten ³¹								
<i>M</i>	4.16	4.17	4.18	4.30	0.01	0.01	0.12	0.14
<i>SE</i>	.032	.045	.073	.085				
B) Ausländer belasten unser soziales Netz								
<i>M</i>	4.06	3.73	3.49	3.41	-0.33	-0.24	-0.08	-0.65
<i>SE</i>	.032	.045	.071	.089				
C) Ausländer bereichern unsere Kultur								
<i>M</i>	4.17	4.46	4.58	4.65	0.29	0.12	0.07	0.48
<i>SE</i>	.032	.044	.069	.085				
D) Ausländer verknappen Wohnungen								
<i>M</i>	2.51	2.37	2.34	2.36	-0.14	-0.03	0.02	-0.15
<i>SE</i>	.029	.040	.065	.083				
E) Ausländer stützen die Rentensicherung								
<i>M</i>	3.75	3.88	4.09	4.24	0.13	0.21	0.15	0.49
<i>SE</i>	.033	.047	.075	.093				
F) Ausländer nehmen Arbeitsplätze weg								
<i>M</i>	3.36	3.03	2.84	2.73	-0.33	-0.19	-0.11	-0.63
<i>SE</i>	.032	.043	.072	.087				
G) Ausländer begehen häufiger Straftaten								
<i>M</i>	4.14	3.91	3.76	3.69	-0.23	-0.15	-0.07	-0.45
<i>SE</i>	.035	.050	.080	.101				
H) Ausländer schaffen Arbeitsplätze								
<i>M</i>	3.16	3.36	3.59	3.64	0.20	0.23	0.05	0.48
<i>SE</i>	.029	.041	.066	.081				

³¹ Vollständiger Wortlaut der Aussagen im Anhang.

Nun könnte der Eindruck entstehen, dass derartige Abweichungen in den Mittelwerten auf einer 7-stufigen Antwortskala nicht übermäßig hoch sind oder zumindest die Items für eine grobe Schätzung im Alltag genügend genau sind. Um dieser Fehleinschätzung bzw. möglichen Fehlinterpretationen vorzubeugen, nachfolgend einige Prozentuierungsvergleiche für die vier Items, mit denen Einstellungen gegenüber in Deutschland lebenden Ausländern erhoben wurden.³²

Hierzu wurden auf der 7-stufigen Antwortskala die Skalenwerte 1-3 zusammengefasst, die eine Ablehnung der Aussagen zum Ausdruck bringen.³³ In Tabelle 4 werden die entsprechenden Prozentwerte der Online-Teilnehmer denen des Gesamt-ALLBUS gegenüber gestellt. Zusätzlich wurde für die Gruppe der Online-Teilnehmer eine Gewichtung vorgenommen, indem deren Struktur nach den Hintergrundmerkmalen Alter, Geschlecht und Schulabschluss an diejenige des Gesamt-ALLBUS angeglichen wurde.

Tabelle 4: Prozentanteil „Ablehner“ (Skalenwerte 1-3) bei 4 Ausländer-Items: Vergleich der Gesamtstichprobe mit den Online-Teilnehmern (gewichtet und ungewichtet)

	ALLBUS- Teilnehmer Gesamt	Online- Teilnehmer (ungewichtet)	Online- Teilnehmer (gewichtet)*	Differenz ALLBUS zu	
				Online ungewichtet	Online gewichtet*
	%	%	%	Prozent- punkte	Prozent- punkte
A) Ausländer: mehr Lebensstilanpassung ³⁴	8,9	13,1	14,1	4,2	5,2
B) Ausländer: wieder heim bei knapper Arbeit	56,4	69,1	65,4	12,7	9,0
C) Ausländer: polit. Betätigung untersagen	55,4	68,8	60,8	13,4	5,4
D) Ausländer: sollten unter sich heiraten	71,2	91,4	90,5	20,2	19,3

* Anpassung an Demographiestruktur der Gesamtstichprobe ALLBUS anhand der Variablen Alter, Geschlecht und Schulabschluss.

Die Zahlen verdeutlichen, dass allein auf der Basis der Online-Teilnehmer keine auch nur halbwegs akzeptable Schätzung der Einstellungen aller ALLBUS-Teilnehmer (der Grundgesamtheit) möglich ist. Die Differenzen belaufen sich auf bis zu 20 Prozentpunkte und auch die Gewichtung führt nicht zu einer entscheidenden Verbesserung.

³² Vgl. hierzu die Items aus Tabelle 2.

³³ Die Ablehnung der Aussagen ist gleichbedeutend mit einer positiven Einstellung gegenüber Ausländern. Vgl. hierzu die Formulierung der einzelnen Aussagen im Anhang.

³⁴ Vollständiger Wortlaut der Aussagen im Anhang

Ein anderes Bild zeigt sich, wenn die Online-Teilnehmer mit der Gesamtheit der Internetnutzer verglichen werden (Tabelle 5). Auch hier erfolgte zusätzlich eine Gewichtung der Online-Teilnehmer, allerdings zur Angleichung an die Struktur aller Internetnutzer. Dies ist plausibel, wenn bei einer Untersuchung alle Internetnutzer die Grundgesamtheit bilden.

Tabelle 5: Prozentanteil „Ablehner“ (Skalenwerte 1-3) bei 4 Ausländer-Items: Vergleich der Internetnutzer mit den Online-Teilnehmern (gewichtet und ungewichtet)

	Internet- Nutzer Gesamt	Online- Teilnehmer (ungewichtet)	Online- Teilnehmer (gewichtet)*	Differenz ALLBUS zu	
				Online ungewichtet	Online gewichtet*
	%	%	%	Prozent- punkte	Prozent- punkte
A) Ausländer: mehr Lebensstilanpassung ³⁵	11,0	13,1	12,8	2,1	1,8
B) Ausländer: wieder heim bei knapper Arbeit	65,9	69,1	66,9	3,2	1,0
C) Ausländer: polit. Betätigung untersagen	62,4	68,8	65,4	6,4	3,0
D) Ausländer: sollten unter sich heiraten	85,1	91,4	91,2	6,3	16,1

* Anpassung an Demographiestruktur der Gesamtheit aller Internetnutzer anhand der Variablen Alter, Geschlecht und Schulabschluss.

Die Differenzen zur Gesamtheit der Internetnutzer bewegen sich in einem deutlich geringeren Umfang. Dies gilt sowohl für die ungewichteten Prozentwerte der Online-Teilnehmer als auch für die gewichteten Daten, die mit Ausnahme des Items „D“ nur noch geringe Abweichungen zur Gesamtheit der Internetnutzer aufweisen. In diesem Ergebnis könnte ein Hinweis darauf gesehen werden, dass mit einer Online-Bevölkerungsumfrage zumindest eine Schätzung für eine Grundgesamtheit möglich ist, die über die (private) Nutzung des Internet zu definieren wäre.

4. Fazit und Ausblick

Ausgangspunkt unserer Studie war die Frage, wo die derzeitigen Grenzen einer online durchgeführten Bevölkerungsumfrage liegen. Anders als z.B. in der Marktforschung oder der experimentell-psychologischen Forschung, in der die Online-Methodik inzwischen eine

³⁵ Vollständiger Wortlaut der Aussagen im Anhang

weite Verbreitung gefunden hat³⁶, unterliegen allgemeine Bevölkerungsumfragen speziellen Anforderungen an die Stichprobenqualität. Entscheidend ist in diesem Zusammenhang die Frage, in welchem Ausmaß (zumindest gegenwärtig) bestimmte Bevölkerungsgruppen aufgrund eines fehlenden Zugangs zum Internet von der Teilnahme an Online-Befragungen ausgeschlossen sind. Die ALLBUS-Erhebung 2006 mit anschließendem Online Follow-Up bot die Möglichkeit, Coverage und Nonresponse Error exakt zu bestimmen.

Auf den ersten Blick geben unsere Ergebnisse diesbezüglich ein ernüchterndes Bild: Nicht einmal die Hälfte der im ALLBUS befragten Personen nutzt privat das Internet – nach wie vor sind es vorwiegend Jüngere mit einem mittleren bzw. hohen Bildungsabschluss. Der hiermit einhergehende Coverage Error zeigt sich beim Mittelwertvergleich substantieller Variablen.

Auf der anderen Seite ist bei gegebener Internetnutzung ein vergleichsweise geringer Nonresponse Error festzustellen, was für die Durchführbarkeit repräsentativer Online-Bevölkerungsumfragen spricht, sofern die Grundgesamtheit über die private Internetnutzung definiert wird. Zu beachten ist jedoch die aufgrund der notwendigen Rekrutierungsstrategie mit 37 Prozent geringe Teilnahmebereitschaft. Es bleibt abzuwarten, ob es sich hierbei aufgrund einer Unsicherheit im Umgang mit dieser neuartigen Befragungsart um ein temporäres Problem handelt oder um ein generelles Phänomen von Online-Umfragen. Bei anderen Rekrutierungsstrategien z.B. mit wiederholten Kontakten und möglichen Incentives sind durchaus höhere Teilnahmebereitschaften zu erreichen.³⁷ Antworten auf diese und ähnliche Fragen sind zu erhoffen aus aktuell anlaufenden Großforschungsvorhaben, wie dem schon eingangs erwähnten MESS-Projekt in den Niederlanden oder hierzulande dem neuen von der *Deutschen Forschungsgemeinschaft* eingerichteten Schwerpunktprogramm „Survey Methodology“.

Literatur

- ADM Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute (2007): Jahresbericht 2006. Frankfurt.
- Bandilla, W. (2003): Die Internetgemeinde als Grundgesamtheit. In: ADM Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute / Arbeitsgemeinschaft Sozialwissenschaftlicher Institute (ASI). Statistisches Bundesamt (Hrsg.): Online-Erhebungen. Sozialwissenschaftlicher Tagungsbericht. Band 7. Bonn, S. 71-83
- Couper, M. P. (2000): Web Surveys. A Review of Issues and Approaches. In: Public Opinion Quarterly, 64, S. 464-494.
- Couper, M. P., Kapteyn, A., Schonlau, M. & Winter, J. (2007): Noncoverage and Nonresponse in an Internet Survey. In: Social Science Research, 36, S. 131-148.
- Dennis, M. J. & Li, R. (2003): Effects of Panel Attrition on Survey Estimates. Paper presented at the Annual Meeting of the American Association for Public Opinion Research.
- Dillman, D. A. (2007): Mail and Internet Surveys – The Tailored Design Method. 2007 Update with New Internet, Visual, and Mixed-Mode Guide. New York.
- Faas, T. (2003): Offline rekrutierte Access Panels: Königsweg der Online-Forschung? In: ZUMA Nachrichten, 53, S. 58-76.
- Frey, J. H., Kunz, G. & Lüschen, G. (1990): Telefonumfragen in der Sozialforschung. Opladen.
- Fries, R. (2006): Großumfragen im World Wide Web. Aachen.

³⁶ Die von der Deutschen Gesellschaft für Online Forschung (DGOF) unterstützte und aktuell im von Halem Verlag erscheinende Buchreihe „Neue Schriften zur Online-Forschung“ gibt über die derzeitigen Entwicklungen einen umfassenden Überblick.

³⁷ Vgl. Scherpenzeel & Das 2007.

- Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E. & Tourangeau, R. (2004): Survey Methodology. New Jersey.
- Lee, S. (2006): Propensity Score Adjustment as a Weighting Scheme for Volunteer Panel Web Surveys. In: Journal of Official Statistics, 22, S. 329-349.
- Medienpädagogischer Forschungsverbund Südwest (2007): JIM-Studie 2007 – Jugend, Information, (Multi-) Media. Stuttgart.
- Müßig-Trapp, P., Bandilla, W. & Weber, S. (2006): Building Up a Student-Convenience-Panel: First Experiences (Recruitment and Registration Procedure). Paper presented at the General Online Research GOR06.
- Reips U.-D. & Bosnjak, M. (2001): Dimensions of Internet Science. Lengerich.
- Scherpenzeel, A. & Das, M. (2007): "True" Internet panels: Evidence from the Netherlands. Paper presented at GESIS-Workshop "Social Research and the Internet: Advances in Applied Methods and New Research Strategies" (Mannheim, November 07).
- Schnell, R., Hill, P. B. & Esser, E. (2005): Methoden der empirischen Sozialforschung. München.
- Schonlau, M., Fricker, R. & Elliott, M. (2002): Conducting Research Surveys Via E-mail and the Web. Santa Monica.
- Van Eimeren, B. & Frees, B. (2007): ARD-Online-Studie 2007 – Internetnutzung zwischen Pragmatismus und YouTube-Euphorie. Media Perspektiven, 8/2007, S. 362-378.
- Wasmer, M., Scholz, E. & Blohm, M. (2007) Konzeption und Durchführung der „Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften“ (ALLBUS) 2006. ZUMA Methodenbericht, 07/09.

Anlage

Wortlaut der in den Tabellen 2, 4 und 5 verwendeten Fragen und Items:

Frage: Bei dieser Frage geht es um die in Deutschland lebenden Ausländer. Auf dieser Liste stehen einige Sätze, die man schon irgendwann einmal gehört hat. Sagen Sie mir bitte zu jedem Satz, inwieweit Sie ihm zustimmen. Der Wert 1 heißt, dass Sie „überhaupt nicht zustimmen“, der Wert 7 heißt, dass Sie „voll und ganz zustimmen“. Mit den Werten dazwischen können Sie Ihre Meinung abstufen.

- A) Die in Deutschland lebenden Ausländer sollten ihren Lebensstil ein bisschen besser an den der Deutschen anpassen.
- B) Wenn Arbeitsplätze knapp werden, sollte man die in Deutschland lebenden Ausländer wieder in ihre Heimat zurückschicken.
- C) Man sollte den in Deutschland lebenden Ausländern jede politische Betätigung in Deutschland untersagen.
- D) Die in Deutschland lebenden Ausländer sollten sich ihre Ehepartner unter ihren eigenen Landsleuten auswählen.

Wortlaut der in den Tabelle 3 verwendeten Fragen und Items:

Frage: Wie ist es mit den folgenden Aussagen über die in Deutschland lebenden Ausländer? Bitte sagen Sie mir anhand der Skala, inwieweit Sie diesen Aussagen zustimmen. 1 = Stimme überhaupt nicht zu; 7 = Stimme voll und ganz zu.

- A) Die in Deutschland lebenden Ausländer machen die Arbeit, die die Deutschen nicht erledigen wollen.
- B) Die in Deutschland lebenden Ausländer sind eine Belastung für das soziale Netz.
- C) Sie sind eine Bereicherung für die Kultur in Deutschland.
- D) Ihre Anwesenheit in Deutschland führt zu Problemen auf dem Wohnungsmarkt.
- E) Sie tragen zur Sicherung der Renten bei.
- F) Sie nehmen den Deutschen Arbeitsplätze weg.
- G) Sie begehen häufiger Straftaten als die Deutschen.
- H) Die in Deutschland lebenden Ausländer schaffen Arbeitsplätze.